(2)

Final Report

Silicon Association Cortex
N00014-90-J-1349

Principal Investigator – Dan Hammerstorm

Submitted by Todd K. Leen
Assistant Professor
tleen@cse.ogi.edu

August 25, 1992

Technical Accomplishments

Under this grant research was conducted at both Adaptive Systems Inc. (ASI) and the Oregon Graduate Institute (OGI). ASI is developing hardware implementations of models derived from studies of olfactory cortex. For these implementations the models were mapped onto the general-purpose neurocomputer developed at ASI. Research at ASI is also directed towards direct silicon implementation of biologically faithful cortex models.

Research at OGI is concerned with the computational capability and theoretical aspects of these models, as well as with modifications that enhance functionality. We applied cortex-inspired models to speech recognition problems and pursued issues of model convergence and computational efficiency.

During the last year of the grant we explored two new lines of inquiry. The first deals with the dynamics of stochastic learning systems, and the second with locally-linear data encoding. These are described in more detail below.

Implementation on High-level Simulator (OGI) The abstract cortex model [1] was implemented in ASI's high-level simulator. This was the first step in mapping the algorithm to ASI's neuro-computer. This implementation provides a template for the direct micro-code implementation on the chip. In addition, this simulator provided us with a vehicle for preliminary studies of the clustering and classification ability of the model.

Implementation in Micro-Code (ASI) We have completed a direct micro-code implementation of the model for the ASI neuro-computer. The code includes embellishments to the algorithm developed by the research group at OGI. Details of the implementation appear in [2].

Computational Capability (OGI) In [3] we have shown that the abstract model described in [1] and implemented on the ASI neurocomputer is a neural implementation of multistage vector quantization [4, 5]. This architecture has a significant storage advantage relative to conventional neural clustering and vector quantization (VQ) algorithms.

Standard neural VQ algorithms require one neuron for each cluster point. That is, each neuron responds to a specific region of the input space. This defines the receptive field of the neuron.

The multi-stage architecture defines cluster points by *combinations* of neural receptive fields. For example, a cortex model with three hierarchical levels of four neurons each generates 64 distinct cluster points using only 12 neurons. A conventional competitive learning algorithm requires 64 neurons to define the same number of cluster points. Our experiments confirm that the cortex model when used for encoding and classifying phonemes provides lower error per neuron than either conventional or tree-structured algorithms [3].

Naturally one does not get something for nothing, and there is a trade-off inherent in the multi-stage structure. Since each neuron in the model participates in defining several clusters, the latter cannot be independently oriented. However, the disadvantage of this constraint appears to be overwhelmed by the advantage gained by the combinatorial efficiency.

Model Dynamics (OGI) We began studies of the dynamics of self-organizing networks under a previous ONR grant and completed that work under the present grant. (Cliff Lau served as science officer for the previous grant, No. N00014-88-K-0329.) We introduced tools from bifurcation theory to treat the dynamics of *learning* in networks with both Hebbian and anti-Hebbian synapses and recurrent lateral connections. Our studies identify the bifurcation types and thus provide *analytic* descriptions of the location of equilibria and limit cycles in the learning behavior.

Our paper describing this work was among fewer than 6 % of the submitted papers chosen for oral presentation at the 1990 Neural Information Processing Systems conference at Denver, Colorado [6]. The work was recently published in *Network : Computation in Neural Systems* [7] and is the subject of invited talks at the 1990 International Society for the Systems Sciences and the 1991 SPIE Conference on Adaptive Signal Processing.

Stochastic Learning Dynamics (OGI) In stochastic learning, parameter (synapse) adjustments are made after each data presentation, rather than after a "batch" of data has been presented. The Lynch-Granger olfactory cortex models are examples of stochastic learning systems. Much of the literature in stochastic learning is concerned with proving convergence to local optima. There is also a body of literature concerned with optimizing the convergence rate to local minima.

Drawing on tools from statistical physics and the theory of stochastic processes, we have developed an analytic approach to study the ensemble properties of stochastic learning. The approach allows us to calculate transient and equilibrium probability distributions. We applied the formalism to study the statistical distribution of times required to pass *between* basins of different local optima, equilibrium distributions, and optimal learning rate schedules.

This work was published in the 1992 IJCNN [8] and presented at the 1992 Snowbird Conference [9]. We have two papers accepted for the 1992 NIPS meeting, one for oral presentation [10], and one for poster presentation [11].

Locally-Linear Encodings (OGI) The second line of inquiry extends the piece-wise constant encoding produced by algorithms like the Lynch-Granger cortex model. In order to achieve more accurate representations, we combine local linear models with vector quantization (VQ). The resulting algorithm produces a non-linear dimension reduction that is superior to both VQ and the globally linear encoding produced by principal component analysis (PCA).

The only competing algorithm for non-linear dimension reduction uses five-layer feed-forward networks trained to perform an identity transformation between the input and output layers. Applied to speech data, our algorithm runs about five times faster than five-layer networks, and produces more accurate encodings.

Publications

Refereed Papers – 6

1. Leen, T.K.: Dynamics of learning in linear feature-discovery networks, *Network: Computation in Neural Systems* **2**, 85, 1991.

2. Leen, T.K.: Weight-Space Dynamics of Recurrent Hebbian Networks, *Advances in Neural Information Processing Systems 3*, Morgan Kauffman, 1991.

3. Leen, T.K., Webb, Max, Rehfuss, S.: Encoding and Classification in a Model of Olfactory Cortex, *International Joint Conference on Neural Networks*, 1991.

4. Means, E. and Hammerstrom, D.: Piriform model execution on a neurocomputer, *International Joint Conference on Neural Networks*, 1991.

5. Leen, T.K. and Orr, G.B.: Weight-Space Probability Densities and Convergence Times for Stochastic Learning, *International Joint Conference on Neural Networks*, Baltimore, June, 1992.

6. Shaudys, F., Leen, T.K.: Feature Selection for Improved Classification, *International Joint Conference on Neural Networks*, Baltimore, June, 1992.

Book Chapters – 2

1. Hammerstrom, D., Leen, T.K., and Means, E.: Dynamics and Implementation of Self-Organizing Networks, in *Advanced Neural Computers*, R. Eckmiller (Ed.), Elsevier Science Publishers B.V. (North-Holland), March, 1990.

2. Hammerstrom, D. and Means, E.: A proposed architecture for a second–generation neurocomputer, in *Olfaction as a Model Systems for Computational Neuroscience*, Joel Davis and Howard Eichenbaum (Eds.), MIT Press, 1991.

Technical Reports and Non-Refereed Papers – 7

1. Leen, T.K.: Weight Dynamics of Recurrent Hebbian Networks, *Proceedings of the 34th Annual conference of the International Society for the Systems Sciences, July, 1990.*

2. Leen, T.K.: Hebbian Learning: Algorithms and Applications, *Proceedings of the 34th Annual conference of the International Society for the Systems Sciences, July, 1990.*

3. Leen, T.K.: Dynamics of Learning in Recurrent Hebbian Networks, *Oregon Graduate Institute Technical Report No. CSE 90-013*, August, 1990.

4. Leen, T.K., Cole, R., Hammerstrom, D., Inouye, J. : Speech Recognition with a Cortex Model : Preliminary Results and Outlook, *Oregon Graduate Institute Technical Report No. CSE 90-022*, June, 1990.

5. Leen, T.K., Webb, Max, Rehfus, S.: Encoding and classification in a model of olfactory cortex, *Oregon Graduate Institute Tech. Rep. CS/E 91-002*, Jan. 1991.

6. Leen, T.K.: A Coordinate-Independent Center Manifold Reduction, *Oregon Graduate Institute Tech. Rep. CS/E 91-023*, December, 1991.

7. Leen, T.K. and Orr, G.B.:Weight-Space Distributions and Convergence Times for Stochastic Learning, Abstract for *Neural Networks for Computing*, Snowbird, Utah, 1992.

Refereed Papers Submitted, Not Yet Published – 3

1. Leen, T.K. and Orr, G.B.: Probability Densities and Basin-Hopping in Stochastic Learning, **accepted** for *Advances in Neural Information Processing Systems 5.* This paper was one of **6.5 %** of the submissions chosen for oral presentation for the 1992 NIPS conference.

2. Leen, T.K. and Moody, J.E.: Probability Densities and Equilibria in Stochastic Learning, **accepted** for *Adances in Neural Information Processing Systems 5*, 1992.

3. Leen, T.K.: A Coordinate-Independent Center Manifold Expansion, submitted to *Physics Letters A.*

Invited Talks – 5

1. Todd K. Leen, "Local Learning in Hebbian Networks, R.S. Dow Neurological Sciences", Institute, Feb. 1991.

2. Todd K. Leen, "Learning in Linear Feature-Discovery Networks", SPIE Conference on Adaptive Signal Processing, July, 1991.

3. Todd K. Leen, "Stochastic Optimization, Neural Networks and the Fokker-Planck Equation", Dept. of Physics, University of Wisconsin, Milwaukee, WI, June, 1992.

4. Todd K. Leen, "Weight-Space Densities and Basin Hopping in Stochastic Learning", NEC Research Laboratory, Princeton, NJ, June, 1992.

5. Dan Hammerstrom, Steve Rehfuss: "Neurocomputing Hardware: Present and Future", The First Swedish National Conference on Connectionism Sept. 1992, Skovde, Sweden.

Patents – 0

Support

- Graduate Students Supported Greater than 25% – 2.
- Post-doc support – 0.
- Female, minority and Asian graduate students and post-docs – 0.

References

[1] Jose Ambros-Ingerson, Richard Granger, and Gary Lynch. Simulation of paleocortex performs hierarchical clustering. *Science*, 247:1344–1348, 1990.

[2] Eric Means and Dan Hammerstrom. Piriform model execution on a neurocomputer. In *Proceedings of the 1991 International Joint Conference on Neural Networks*, June 1991, submitted.

[3] Todd K. Leen, Max Webb, and Steven Rehfuss. Encoding and classification in a model of olfactory cortex. In *International Joint Conference on Neural Networks*, July 1991.

[4] Biing-Hwang Juang and A.H. Gray Jr. Multiple stage vector quantization for speech coding. In *Proceeding of the IEEE International Conference on Acoustics and Signal Processing*, pages 597–600, 1982.

[5] Robert M. Gray. Vector quantization. *IEEE ASSP Magazine*, pages 4–29, April 1984.

[6] Todd K. Leen. Dynamics of learning in recurrent feature-discovery networks. In Richard P. Lippmann, John Moody, and David Touretzky, editors, *Advances in Neural Information Processing Systems 3*. Morgan Kauffmann, 1991.

[7] Todd K. Leen. Dynamics of learning in linear feature-discovery networks. *Network : Computation in Neural Systems*, 2:85–105, 1991.

[8] Todd K. Leen and Genevieve B. Orr. Weight-space probability densities and convergence times for stochastic learning. In *International Joint Conference on Neural Networks*, pages IV 158–164. IEEE, June 1992.

[9] Todd K. Leen and Genevieve B. Orr. Weight-space distributions and convergence times for stochastic learning. Abstract for Neural Networks for Computing Converence, Snowbird, Utah, 1992.

[10] Todd K. Leen and Genevieve B. Orr. Probability densities and basin hopping in stochastic learning. Accepted for NIPS 92, 1992.

[11] Todd K. Leen and John Moody. Probability densities and equilibria in stochastic learning. Accepted for NIPS 92, 1992.